

Nikunj C. Oza, Ph.D.

NASA Ames Research Center
Mail Stop 269-3
Moffett Field, CA 94035-1000
E-mail: oz@email.arc.nasa.gov
<http://ic.arc.nasa.gov/~oza>
Tel: (650)604-2978
Fax: (650)604-4036

EDUCATION

- Ph.D. **University of California, Berkeley**
Computer Science, September 2001.
Thesis: **Online Ensemble Learning**
- M.S. **University of California, Berkeley**
Computer Science, May 1998.
Thesis: **Probabilistic Models of Driver Behavior**
- B.S. **Massachusetts Institute of Technology**
Mathematics with Computer Science, February 1994.

BOOK CHAPTER (refereed)

- Nikunj C. Oza, Ensemble Data Mining Methods, *(to appear) Encyclopedia of Data Warehousing and Mining*, April 2005.

JOURNAL PUBLICATIONS

- Ashok N. Srivastava, Nikunj C. Oza, and Julianne Stroeve, Virtual Sensors: Using Data Mining to Efficiently Estimate Spectra, *(to appear) IEEE Transactions on Geosciences and Remote Sensing, Special Issue on Advances in Techniques for Analysis of Remotely Sensed Data*, 2004.
- Kagan Tumer and Nikunj C. Oza, Input Decimated Ensembles, *Pattern Analysis and Applications*, 6(1):65-77, 2003.

REFEREED CONFERENCE PUBLICATIONS

- Nikunj C. Oza, AveBoost2: Boosting for Noisy Data, *Fifth International Workshop on Multiple Classifier Systems*, Cagliari, Italy. 2004.
- Ashok N. Srivastava, Nikunj C. Oza, and Julianne Stroeve, Virtual Sensors: Using Data Mining to Efficiently Estimate Spectra, *submitted to the International Geoscience and Remote Sensing Symposium*, 2004.
- Nikunj C. Oza, Boosting with Averaged Weight Vectors, *Fourth International Workshop on Multiple Classifier Systems*, Guildford, UK. 2003.

- Nikunj C. Oza, Kagan Tumer, Irem Y. Tumer, and Edward M. Huff, Classification of Aircraft Maneuvers for Fault Detection, *Fourth International Workshop on Multiple Classifier Systems*, Guildford, UK. 2003.
- Nikunj C. Oza and Stuart Russell, Experimental Comparisons of Online and Batch Versions of Bagging and Boosting, *The Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California. 2001.
- Nikunj C. Oza and Kagan Tumer, Input Decimation Ensembles: Decorrelation through Dimensionality Reduction, *Second International Workshop on Multiple Classifier Systems*, Cambridge, UK. 2001.
- Nikunj C. Oza and Stuart Russell, Online Bagging and Boosting, *Eighth International Workshop on Artificial Intelligence and Statistics*, Key West, Florida. 2001.
- Kagan Tumer and Nikunj C. Oza, Decimated Input Ensembles for Improved Generalization, *International Joint Conference on Neural Networks*, Washington, D.C. 1999, **Recipient: Best Presentation Award**.
- Alex Cuthbert, Christopher Stecker, Inna Aleksandrovsky, Sheryl Ehrlich, Nikunj Oza, and Paula Rogers, Instructional Effects on Spatial and Temporal Memory for Videotaped Events in a Large-scale Environment, *Nineteenth Annual Conference of the Cognitive Science Society*, Stanford, CA. 1997.

OTHER PUBLICATIONS

- Ashok N. Srivastava and Nikunj C. Oza, Knowledge Driven Image Mining with Mixture Density Mercer Kernels, *ESA-EUSC 2004: Workshop on the Theory and Applications of Knowledge Driven Image Information Mining*, with focus on Earth Observation, 2004.
- Ashok Srivastava, Julianne C. Stroeve, and Nikunj C. Oza, Using Kernel Methods to Detect Clouds, Snow, Ice and other Geophysical Processes, *Transactions of the American Geophysical Union*, 84(46), Fall Meeting Supplement, Abstract C12A08-65, 2003.
- Nikunj C. Oza, Online Ensemble Learning, Ph.D. thesis, University of California, Berkeley, 2001.
- Nikunj C. Oza, Online Ensemble Learning, *17th National Conference on Artificial Intelligence, Doctoral Consortium*, Austin, TX. 2000.
- Nikunj C. Oza and Kagan Tumer, Dimensionality Reduction through Classifier Ensembles, Technical Report NASA-ARC-IC-1999-124.
- Nikunj C. Oza, Probabilistic Models of Driver Behavior, *Spatial Cognition Conference*, Berkeley, CA. 1999.
- Nikunj C. Oza, Probabilistic Models of Driver Behavior, M.S. report, University of California, Berkeley, 1998.

Jeffrey Forbes, Nikunj Oza, Ronald Parr, Stuart Russell, Feasibility Study of Fully Automated Vehicles Using Decision-Theoretic Control, California PATH Research Report UCB-ITS-PRR-97-18.

HONORS AND AWARDS

- NASA Incentive Award (for collaboration with Puresense Environmental, NASA Research Park), 2003.
- Best Presentation Award, International Joint Conference on Neural Networks, Washington, D.C., 1999
- Schlumberger Foundation Fellowship, 1999
- NSF Graduate Research Training Program in Cognitive Science (Fellowship), 1996-1998
- California Fellowship in Microelectronics, 1994-1995
- MIT Class of 1950 Scholarship 1990-1993
- Life Membership, California Scholarship Federation, 1990

PROFESSIONAL ACTIVITIES

Conference Chair, Sixth International Workshop on Multiple Classifier Systems, June 2005.

Member of Program Committee, Workshop on Ensembles in Extreme Environments, IEEE Conference on Systems, Man, and Cybernetics. October, 2005.

Member of Program Committee, Conference on Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security; SPIE Defense and Security Symposium, March 2005.

Reviewed articles for

- IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)
- Pattern Recognition Letters
- Encyclopedia of Data Warehousing and Mining
- International Workshop on Multiple Classifier Systems
- Information Fusion (Journal)
- Journal of Machine Learning Research,
- IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control
- International Journal for Pattern Recognition and Artificial Intelligence,
- SIAM Workshop on Data Mining and Machine Learning for Counter-Terrorism
- SIAM Workshop on High Performance Data Mining
- Experimental Program to Stimulate Competitive Research

Invited talk, AAAI National Conference on Artificial Intelligence Doctoral Consortium, July 2004.

Invited Presentation, Information Systems Audit and Control Association (ISACA) San Jose State University chapter meeting, April 2004.

NASA Grant Writing Seminar, September 10-11, 2003.

NASA Project Management course, Jan 27, 2003—Feb 6, 2003.

Invited Presentation, “Machine Learning and Data Mining in System Health and Safety,” California Menay Institute, July 1, 2002.

RESEARCH EXPERIENCE

- Oct. 2004
to
Present *Program Manager, Knowledge Discovery Tools for System-Wide Security, NASA Ames Research Center, Moffett Field, California.*
Determined research plan for data mining applied to aviation security problems, including flight deviation and cargo screening.
- Sep. 2001
to
Present *Research Scientist, NASA Ames Research Center, Moffett Field, California.*
Worked on fault detection for aircraft. Currently working on ensemble learning methods (input decimation, bagging, boosting), applications of machine learning to satellite image understanding, water distribution system monitoring, and aviation security.
- Jun. 1998—
Aug. 1998
and
Jun. 1999—
Aug. 1999 *Summer Researcher, NASA Ames Research Center, Moffett Field, California.*
Worked on input decimation--a new method of correlation reduction in ensemble classifiers that presents different features to different individual classifiers within the ensemble. Achieved classification results several standard deviations above the level of standard neural networks on several well-known datasets within the UCI Machine Learning Repository and numerous synthetic datasets.
- Aug. 1996
to
May 1998 *Member, NSF Graduate Research Training Program in Spatial Cognition, Institute for Cognitive Studies, University of California, Berkeley.*
Worked with Professor Stephen Palmer and students from several departments on the issue of separability of spatial and sequential mental representations as revealed by having subjects view a video depicting movement through a large-scale space and having them draw maps and make ordered lists of events occurring in the video.
- Oct. 1995
to
Aug. 2001 *Graduate Student Researcher, Computer Science Division, University of California, Berkeley.*
Worked with Professor Stuart Russell on Bayesian Automated Taxi project.
Worked on learning probabilistic models of automobile driver behavior to allow an automated vehicle to predict the actions of nearby vehicles, thereby improving its own driving. Designed online versions of the bagging and boosting ensemble learning algorithms.
- Jun. 1995
to
Aug. 1995 *Researcher, Information Technology Laboratory, General Electric Corporate Research and Development, Schenectady, New York.*
Worked with Dr. Tomek Strzalkowski on the Natural Language Toolkit, a C++ library of tools to analyze documents. Specifically, designed and implemented module to classify sentences as imperative, declarative, fragments, or questions. Used this module to find procedures---sequences of imperative sentences---in maintenance manuals.

TEACHING EXPERIENCE

- Jan. 1995 *Teaching Assistant, University of California, Berkeley*
to
Dec. 1999 Structure and Interpretation of Computer Programs (Fall 1999), Artificial
Intelligence (Fall 1995, Fall 1998), Computer Architecture (Spring 1995).
- May 1999 *Member, Part-time Instructor Pool, City College of San Francisco,*
to
May 2000 *California.*
Awarded membership in part-time pool, allowing members to choose each
semester whether to teach and which courses to teach each semester.
- Jan. 1998 *Instructor, City College of San Francisco, California.*
to
May 1998 Designed and lectured intermediate C++ course.
- Jan. 1993 *Instructor, MIT Educational Studies Program, Cambridge, Massachusetts*
to
May 1993 Designed and taught a course in Artificial Intelligence at MIT.
- Sep. 1990 *Instructor, MIT Educational Studies Program, Cambridge, Massachusetts*
to
Nov. 1991 Taught courses on the Mathematics portion of the Scholastic Aptitude Test.
(SAT).

OTHER PROFESSIONAL EXPERIENCE

- Nov. 1996 *Independent Item Writer, Educational Testing Service, Oakland, California*
to
Nov. 1998 Independent author of questions for the new Graduate Record Examination (GRE)
Mathematical Reasoning (MR) test.
- Jan. 1994 *Engineer, General Electric Nuclear Energy, San Jose, California*
to
Aug. 1994 Evaluated several potential document management solutions for GE Nuclear
Energy. Performed Database and Document Management System Administrator
duties for Verity TOPIC databases (required C, UNIX, shell script, and MS-DOS
batch programming). Established network remote access for Materials Services
Site Representatives.
- Jun. 1993 *Summer Intern, General Electric Nuclear Energy, San Jose, California*
to
Aug. 1993 Tested and wrote operator's manual for General Electric's ultrasonic Remote
Inspection System and Motion Control System (a four-axis motor control system).
- Jun. 1992 *Summer Intern, General Electric Nuclear Energy, San Jose, California*
to
Sep. 1992 Set up data acquisition and analysis system for fuel rod vibration data analysis
using Viewdac software. Established it as part of standard operating procedure for
vibration testing by performing required testing and writing appropriate
documentation.
- Apr. 1992 *Chairman, MIT Educational Studies Program, Cambridge, Massachusetts*
to
to Made all major executive decisions and oversaw the work of the directors of each

May 1993 educational program.

Jan. 1992 *Director, SAT Preparation Program, MIT Educational Studies*
to *Program, Cambridge, Massachusetts*

Apr. 1992 Directed program of 16 teachers and 150 students. Made all executive decisions specifically involving the SAT Preparation Program, presented reports on this program as a member of the Executive Board, and provided support to teachers.

May 1991 *Researcher, Undergraduate Research Opportunity Program, Massachusetts*
to *Institute of Technology, Cambridge, Massachusetts*

Sep. 1991 Designed and wrote a Survival Analysis software package for Professor Michael Rappa of the Sloan School of Management at MIT. Wrote the package in the C programming language for the Apple Macintosh series computers. Designed and implemented the user interface, algorithms, and code to perform survival data analysis.

PROFESSIONAL MEMBERSHIPS

American Association for Artificial Intelligence (AAAI)
Association for Computing Machinery (ACM)

PROFESSIONAL MEMBERSHIPS

United States citizen (by birth)
Interim Secret clearance